AMENDMENTS TO THE CLAIMS

1. (Currently amended) At least one computer readable medium encoded with instructions that, when executed by at least one processor, perform a method for generating a speech recognition model ~~models~~, the method comprising:

receiving [[a]] female speech training data;

generating female ~~recognition model of~~ phoneme models based on [[a]] the female ~~set of recorded phonemes~~ speech training data;

receiving a male speech training data;

generating male ~~recognition model of~~ phoneme models based on [[a]] the male ~~set of recorded phonemes~~ speech training data;

determining a difference ~~in model information~~ between ~~pairs of~~ each female phoneme model and each corresponding male phoneme model ~~models of the female speech recognition model and the male speech recognition model~~; [[and]]

creating a gender-independent phoneme model ~~speech recognition model that includes a gender-independent phoneme model based on a pair of corresponding phoneme models of the female speech recognition model and the male speech recognition model~~ when the difference between the compared female phoneme model and the corresponding male ~~in model information between the phoneme models of the pair of corresponding~~ phoneme model ~~models~~ is less than a predetermined value ~~insignificant~~; and

adding, based on at least one criteria, one of the gender-independent phoneme model, or both the female phoneme model and the corresponding male phoneme model to the speech recognition model.

2. (Currently amended) The at least one computer readable medium of claim 1, ~~further comprising removing each of the phoneme models of the pair of corresponding phoneme models from the female speech recognition model and the male speech recognition model when the difference in model information between the phoneme models is insignificant~~ wherein the at least one criteria comprises a threshold value or an upper limit for the total number of phoneme models in the speech recognition model.

3. (Currently amended) The at least one computer readable medium of claim 1, wherein determining the difference ~~in model information~~ includes calculating a Kullback Leibler distance between the ~~first speech recognition~~ each female phoneme model and ~~second speech recognition~~ the each corresponding male phoneme model.

4. (Currently amended) The at least one computer readable medium of claim 3, wherein ~~whether the model information is insignificant~~ the difference is ~~based on~~ a ~~threshold~~ Kullback Leibler distance quantity.

5. (Currently amended) The at least one computer readable medium of claim 1, wherein the female ~~speech recognition~~ phoneme models ~~model~~, the male ~~speech recognition~~ phoneme models ~~model~~, and the gender-independent ~~speech recognition~~ phoneme models ~~model~~ are Gaussian mixture models.

6. (Currently amended) A system for generating a speech recognition model ~~models~~, the system comprising:

an input to receive speech training data; and

a computer processor coupled to the input, the computer processor configured to: [[;]]

receive a first ~~speech recognition model of phoneme models based on a first~~ set of speech training data, the first set of speech training data originating from a first set of common entities;

generate first phoneme models based on the first set of speech training data;

receive a second ~~speech recognition model of phoneme models based on a second~~ set of speech training data, the second set of speech training data originating from a second set of common entities; [[and]]

generate second phoneme models based on the second set of speech training data;

determine a difference between each first phoneme model and each corresponding second phoneme model;

~~a processing module configured to~~ create an independent ~~speech recognition~~ phoneme model ~~that includes an independent phoneme model based on a pair of corresponding phoneme models of the first speech recognition model and the second speech recognition model~~ when the difference ~~in model information~~ between the compared each first phoneme model ~~models~~ and each ~~of the pair of~~ corresponding second phoneme model ~~models~~ is less than a predetermined value ~~insignificant~~; and

add, based upon at least one criteria, one of the independent phoneme model, or both the first phoneme model and the corresponding second phoneme model to the speech recognition model.

7. (Currently amended) The system of claim 6, wherein the at least one criteria comprises a threshold value or an upper limit for the total number of phoneme models in the speech recognition model. ~~processing module is configured to remove each of the phoneme models of the pair of corresponding phoneme models from the first speech recognition model and the second speech recognition mode when the difference in model information between the phoneme models is insignificant~~.

8. (Currently amended) The system of claim 6, wherein the ~~processing model~~ computer processor is further configured to calculate a Kullback Leibler distance between the each first phoneme ~~speech recognition~~ model and the each corresponding second ~~speech recognition~~ phoneme model.

9. (Currently amended) The system of claim 8, wherein ~~whether~~ the difference ~~model information is insignificant~~ is ~~based on~~ a ~~threshold~~ Kullback Leibler distance quantity.

10. (Currently amended) The system of claim 6, wherein the first ~~speech recognition~~ phoneme models ~~model~~, the second ~~speech recognition~~ phoneme models ~~model~~, and the independent ~~speech recognition~~ phoneme models ~~model~~ are Gaussian mixture models.

11. (Currently amended) A computer program product embodied in computer memory comprising:

computer readable program codes ~~coupled to the~~ executable on a computer ~~memory~~ system for generating a speech recognition model ~~models~~, the computer readable program codes configured to cause the ~~program~~ system to:

receive a first ~~speech recognition model of phoneme models based on a first~~ set of speech training data, the first set of speech training data originating from a first set of common entities;

generate first phoneme models based on the first set of speech training data;

receive a second ~~speech recognition model of phoneme models based on a second~~ set of speech training data, the second set of speech training data originating from a second set of common entities; [[and]]

generate second phoneme models based on the second set of speech training data;

determine a difference ~~in model information~~ between ~~pairs of corresponding phoneme models of the~~ each first ~~speech recognition~~ phoneme model and ~~the~~ each second ~~speech recognition~~ phoneme model; ~~and~~

create an independent ~~speech recognition~~ phoneme model ~~that includes an independent phoneme model based on a pair of corresponding phoneme models of the first speech recognition model and the second speech recognition model~~ when the difference ~~in model information~~ between the each first phoneme model ~~models of the pair of~~ and the each corresponding second phoneme model ~~models~~ is less than a predetermined value ~~insignificant~~; and

add, based on at least one criteria, one of the independent phoneme model, or both the first phoneme model and the corresponding second phoneme model to the speech recognition model.

12. (Currently amended) The computer program product of claim 11, wherein the at least one criteria comprises a threshold value or an upper limit for the total number of phoneme models in the speech recognition model ~~the computer readable program codes configured to cause the program to remove each of the phoneme models of the pair of corresponding phoneme models from the first speech recognition model and the second speech recognition model when the difference in model information between the phoneme models is insignificant~~.

13. (Currently amended) The computer program product of claim 11, wherein the determining the difference ~~in model information~~ includes calculating a Kullback Leibler distance between the each first phoneme model and the each corresponding second phoneme model.

14. (Currently amended) The computer program product of claim 13, wherein ~~whether the model information is insignificant~~ the difference is ~~based on~~ a threshold Kullback Leibler distance quantity.

15. (Currently amended) The computer program product of claim 11, wherein the first ~~speech recognition~~ phoneme models ~~model~~, the second ~~speech recognition~~ phoneme models ~~model~~, and the independent ~~speech recognition~~ phoneme models ~~model~~ are Gaussian mixture models.

16. (Cancelled)

17. (Currently amended) At least one computer readable medium encoded with instructions that, when executed by at least one processor, perform a method for recognizing speech from an audio stream originating from one of a plurality of data classes, each data class having class-dependent phoneme models, the method comprising:

receiving a current feature vector of the audio stream;

computing ~~a current vector probability~~ best estimates that the current feature vector belongs to each one of the plurality of data classes;

computing [[an]] accumulated confidence values for each of the plurality of data classes ~~level that the audio stream~~ that the current feature vector belongs to each one of the plurality of data classes, the confidence value for each data class of the plurality of data classes based on the current ~~vector probability~~ best estimate for the data class and on previous confidence values for the data class, the previous confidence values associated with previous feature vectors of the audio stream ~~vector probabilities~~;

weighing the class-dependent phoneme models based on the accumulated confidence values; and

recognizing the current feature vector based on the weighted class-dependent phoneme models~~; and~~

~~wherein the plurality of data classes include a female speech recognition model based on recorded phonemes originating from a plurality of female speakers, a male speech recognition model based on recorded phonemes originating from a plurality of male speakers, and a gender-independent speech recognition model that includes independent phoneme models based on pairs of corresponding recorded phonemes originating from the plurality of female speakers and the plurality of male speakers determined to have insignificant differences in model information between the recorded phonemes of the pair of corresponding recorded phonemes, each of the female speech recognition model and the male speech recognition model lacking the phoneme models of the gender-independent speech recognition model based on pairs of corresponding recorded phonemes originating from the plurality of female speakers and the plurality of male speakers determined to have insignificant differences in model information between the recorded phonemes of pairs of corresponding recorded phonemes~~.

18. (Currently amended) The at least one computer readable medium of claim 17, wherein computing ~~the current vector probability~~ best estimates includes estimating an a posteriori class probability for the current feature vector.

19. (Currently amended) The at least one computer readable medium of claim 17, wherein computing [[the]] accumulated confidence ~~level~~ values further ~~comprising~~ comprises weighing the current ~~vector probability~~ confidence values more than the previous ~~vector probabilities~~ confidence values.

20. (Previously presented) The at least one computer readable medium of claim 17, the method further comprising determining if another feature vector is available for analysis.

21. (Currently amended) A system for recognizing speech data from an audio stream originating from one of a plurality of data classes, each data class having class-dependent phoneme models, the system comprising:

a computer processor;

a receiving module configured to receive a current feature vector of the audio stream;

a first computing module configured to compute [[a]] current ~~vector probability~~ best estimates that the current feature vector belongs to each one of the plurality of data classes;

a second computing module configured to compute [[an]] accumulated confidence ~~level~~ values for each of the plurality of data classes that the ~~audio stream~~ current feature vector belongs to each one of the plurality of data classes, the confidence value for each data class of the plurality of data classes based on the current ~~vector probability~~ best estimate for the data class and on previous ~~vector probabilities~~ confidence values for the data class, the previous confidence values associated with previous feature vectors of the audio stream;

a weighing module configured to weigh the class-dependent phoneme models based on the accumulated confidence values; and

a recognizing module configured to recognize the current feature vector based on the weighted class-dependent phoneme models~~; and~~

~~wherein the plurality of data classes include a first speech recognition model based on recorded phonemes originating from a first set of speakers, a second speech recognition model based on recorded phonemes from a second set of speakers, and a third speech recognition model that includes phoneme models based on pairs of corresponding recorded phonemes originating from both the first and second set of speakers determined to have insignificant differences in model information between the recorded phonemes of the pair of corresponding recorded phonemes, each of the first speech recognition model and the second speech recognition model lacking the phoneme models of the third speech recognition model based on pairs of corresponding recorded phonemes originating from both the first and second set of speakers determined to have insignificant differences in model information between the recorded phonemes of the pairs of corresponding recorded phonemes~~.

22. (Original) The system of claim 21, wherein the first computing module is further configured to estimate an a posteriori class probability for the current feature vector.

23. (Currently amended) The system of claim 21, wherein the second computing module is further configured to weigh the current ~~vector probability~~ confidence values more than the previous ~~vector probabilities~~ confidence values.

24. (Currently amended) A computer program product embodied in computer memory comprising:

computer readable program codes ~~coupled to the~~ executable on a computer ~~memory~~ system for recognizing speech data from an audio stream originating from one of a plurality of data classes, each data class having class-dependent phoneme models, the computer readable program codes configured to cause the ~~program~~ system to:

receive a current feature vector of the audio stream;

compute ~~a current vector probability~~ best estimates that the current feature vector belongs to each one of the plurality of data classes;

compute [[an]] accumulated confidence values for each of the plurality of data classes ~~level that the audio stream~~ that the current feature vector belongs to each one of the plurality of data classes, the confidence value for each data class of the plurality of data classes based on the current ~~vector probability~~ best estimate for the data class and on previous confidence values for the data class, the previous confidence values associated with previous feature vectors of the audio stream ~~vector probabilities~~;

weigh the class-dependent phoneme models based on the accumulated confidence values; and

recognize the current feature vector based on the weighted class-dependent phoneme models~~; and~~

~~wherein the plurality of data classes include a first speech recognition model based on recorded phonemes originating from a first set of speakers, a second speech recognition model based on recorded phonemes from a second set of speakers, and a third speech recognition model~~

~~that includes phoneme models based on pairs of corresponding recorded phonemes originating from both the first and second set of speakers determined to have insignificant differences in model information between the recorded phonemes of the pairs of corresponding recorded phonemes, each of the first speech recognition model and the second speech recognition model lacking the phoneme models of the third speech recognition model based on pairs of corresponding recorded phonemes originating from both the first and second set of speakers determined to have insignificant differences in model information between the recorded phonemes of the pairs of corresponding recorded phonemes~~.

25. (Currently amended) The computer program product of claim 24, wherein the program code configured to cause the system to compute the current ~~vector probability~~ best estimates includes program code configured to cause the system to determine an a posteriori class probability for the current feature vector.

26. (Currently amended) The computer program product of claim 24, wherein the program code configured to cause the system to compute the accumulated confidence ~~level~~ values includes program code configured to cause the system to weigh the current ~~vector probability~~ confidence values more than the previous ~~vector probabilities~~ confidence values.

27. (Currently amended) The computer program product of claim 24, further comprising program code configured to cause the system to determine if another feature vector is available for analysis.